

WHAT IS CLAIMED IS:

1. A method for categorizing a set of objects, comprising:
 - defining a set of categories in which at least one category in the set is dependent on another category in the set;
 - organizing the set of categories in a hierarchy that embodies any dependencies among the categories in the set;
 - for each object, assigning to the object one or more categories $l_1 \dots l_p$, where $l_i \in \{1 \dots L\}$ from a set $\{1 \dots L\}$ of possible categories, wherein the assigned categories represent a subset of categories for which the object is relevant;
 - defining a new set of labels z comprising all possible combinations of any number of the categories, $z \in \{\{1\}, \{2\}, \dots, \{L\}, \{1,2\}, \dots, \{1,L\}, \{2,3\}, \dots, \{1,2,3\}, \dots, \{1,2, \dots, L\}\}$, such that if an object is relevant to several categories, the object must be assigned the unique label z corresponding to the subset of all relevant categories; and
 - assigning to the object the several categories and the subcategories of the several categories.
2. The method of claim 1, wherein an object comprises a document d generated by co-occurrence of words within the document.
3. The method of claim 2, wherein the hierarchy is generated by:
 - for each document d , choosing a document category α according to the probability $P(\alpha | d) \propto P(d | \alpha) P(\alpha)$;
 - selecting a label v according to the category-conditional probability $P(v | \alpha)$;
 - selecting a word in the document according to a label-specific word distribution $P(w | v)$; and
 - restricting $P(v | \alpha)$ to give positive probability only to labels that are above the category in the hierarchy.

4. The method of claim 1, wherein an object is expressed as a vector in an arbitrary vector space.

5. The method of claim 1, wherein the set of categories further includes an independent category for including objects that fit none of the other categories in the set.

6. The method of claim 1, wherein the hierarchy comprises a directed acyclic graph.

7. The method of claim 1, wherein the hierarchy organizes the sets of categories which embodies dependencies between the categories.

8. The method of claim 3, wherein the category-conditional probability is defined as a uniform distribution over all nodes above the first class.

9. The method of claim 3, wherein the category-conditional probability is defined such that the probability is inversely proportional to a distance of a node to a first class.

10. The method of claim 3, wherein the category-conditional probability of a node is proportional to the number of documents that belong to the node.